DOCUMENT RESUME

ED 213 406 IR 010 033

TITLE Indexation Rules for Metallurgy in PASCAL. Original Title: Regles d'Indexation de la Metallurgie'--Technical Note Issued by Informascience--January 1980. Translated by Marie Wallin.

INSTITUTION Royal Inst. of Tech., Stockholm (Sweden). Library.

REPORT NO TRITA-LIB-4070

PUB DATE Sep 81

NOTE 14p.; For related document, see ED 185 974.

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Databases; French; German; Guidelines; *Indexing; Information Retrieval; International Programs; *Metallurgy; Online Systems; Research Reports; *Subject Index Terms

IDENTIFIERS France; *PASCAL Data Base; West Germany

ABSTRACT

The indexing rules presented are designed for use with a new French-German database on metallurgy being developed under an agreement by CNRS (Centre National de la Recherche Scientifique, Paris) and BAM (Bundesanstalt fur Materialprufung, Berlin). The new database, which will feature multilingual titles and index terms (French-German-English-) and bilingual abstracts (French-German), is expected to acquire and process some 30,000 references per year. The indexing rules given reflect some changes in the 1979 rules for the metallurgy section of PASCAL, primarily in the classification codes, and in the indexing and selection of input documents. Types of descriptors and other subject terms employed in the system are listed and described, the chaining together of controlled vocabulary and free text terms in the PASCAL system is explained, and the use of supplementary classification codes is outlined. An appendix on the application of PASCAL in the field of metallurgy using the EPOS/VIRA computer system is attached. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED213406

REPORT TRITA-LIB-4070 SEPTEMBER 1981

INDEXATION RULES FOR METALLURGY IN PASCAL

ORIGINAL TITLE: REGLES D'INDEXATION DE LA METALLURGIE' - TECHNICAL NOTE ISSUED BY INFORMASCIENCE - JAN. 1980

translated by Marie Wallin

Kungliga Tekniska Högskolans Bibliotek
100 44 STOCKHOLM

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Stephen Schwarz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

INDEXATION RULES FOR METALLURGY IN PASCAL

CNRS (Centre National de la Recherche Scientifique, Paris) and BAM (Bundesanstalt für Materialprüfung, Berlin) under the sponsorship of the French MIDIST (Mission Interministérielle de l'Information Scientifique et Technique) and BMFT (Bundesministerium für Forschung und Technologie, Bonn) have concluded an agreement on the building of a common French-German data base in the field of metallurgy, starting 1980.

This data base will be
- multilingual (French- German- English) for titles and indexation
- bilingual (French- German) for abstracts.

This implies some changes in the section 740 (Metallurgy) of PASCAL (and Bulletin Signalétique) when compared to what was the case in 1979. Changes concern mainly classification codes (see Plan de Classement 1980 and later) the indexation and the selection of input documents (the data base will comprehend some 30.000 references a year).

Practical details on the indexing rules are given below. The same indexing rules apply to section 745 of PASCAL: Welding.

1 DESCRIPTORS

Four descriptor types are used: controlled descriptors
identifiers
candidate descriptors
free terms

They are registered in the following order:
French controlled descriptors
identifiers
candidate descriptors
English controlled descriptors
German controlled descriptors
free terms

1.1 Controlled descriptors

These are the descriptors found in the controlled word list (Lexique) for Metallurgy, based on the multilingual Metallurgy Thesaurus SDIM Thesaurus that will be published by C.E.C (a pre-print version was issued Oct. 1, 1979). The Lexique for Metallurgy is now (1981) published by Informascience in the trilingual form French- English, French- German and is computer stored so that automatical translation is possible between all of these terms. The list includes some 600 more terms than Metallurgy Thesaurus mainly in the field of welding. The same syntaxic rules are applied to the French terms as in the other PASCAL Lexiques.

Ex: ACIER INOXYDABLE DURISSEMENT ....
STAINLESS STEEL HARDENING ....
NICHTROSTENDER STAHL HAERTEN ....

1.2 Identifiers

The identifiers are by definition not included in the list of controlled words (Lexique), they represent materials.
The identifiers are written according to rules which can be considered more or less as international as they are using chemical symbols and formula.

In general: materials are given in the form of chemical symbols or as chemical formula of the constituents

two constituents are separated by a space or by two numbers representing the standard range (in percent) of their composition (see example below)

the composition range, when given, is separated from the chemical symbol concerned by a ':'

Ex: MN:0,1-0,5 SI:0,1-0,5 C:0-0,23 (note the spaces!)

the different constituents of intermetallic compounds are given in the alphabetical order of their symbols

Ex: $Pt_{13}In_9$ is written IN9PT13

the other chemical compounds (inclusive nitrides, carbides, borides, etc....) are written in the order cation-anion

Ex: NACL WC FES

the indices for the intermetallic and chemical compounds are written in line with the symbols

Ex: Al_2O_3 is written AL203

Prefixes have been introduced to give the best possible precision in the description of the materials. These are written in front of the symbol of the material and are separated from it by a space. The used prefixes are:

ACIER	steel without indication of composition
ACIERA	alloy steel
ACIERNA	non alloy steel
ALLIAGE	alloy with unknown number of elements
ALLIAGE2	bi-alloy
ALLIAGE3	tri-alloy
ALLIAGEM	alloy with more than three elements
ALLIAGEX	reinforced alloys by non-metallic diffusion, cermet
METAL PUR	metal
FONTE	cast iron without indication of composition
SYSTEME2	binary system
SYSTEME3	ternary system
SYSTEME4	quaternary system
SYSTEMEM	system with more than four components

N.B. there are no spaces inside of the different prefixes except for pure metal which is written in two words: METAL PUR.

Prefixes can also be used alone. They are compulsory in the indexation of systems, steels, cast iron, alloys and dispersed phase compounds. No prefixes are used for the description of intermetallic compounds.

For each of the categories corresponding to these prefixes the order in which the different elements are given is fixed as follows:

- Alloys: the basic element, if there is such a one, is named first. The other elements are written thereafter in the alphabetical order of their chemical symbol. Additions and impurities are named last, also in alphabetical order.

  Ex: ALLIAGEX AG:>50 CDO:10-20
  ALLIAGEX CO:5-10 TIN:10-15 WC:>50
  ALLIAGEX CU:>50 BEO

  n.b. the sign > (greater than) can appear in composition range instead of a number.

- Dispersion strengthened composites: the matrix is named first and the dispersed phases last even if their proportion is larger than that of the matrix.

  Ex: ALLIAGEX AG:>50 CDO:10-20
  ALLIAGEX CO:5-10 TIN:10-15 WC:>50

- Steel and cast iron: the iron is never named. The other components are all named in the alphabetical order followed by C (carbon symbol). Impurities and additions are given after carbon, themselves in alphabetical order. Distinction between alloy steels and non-alloy steels is made according to the Eurostandard 20-74.

  Ex: ACIERNA C:0-0,23 Nb:0,01-0,05

- Systems: all the constituents are named in the alphabetical order.

Ex: SYSTEME3 CR FE NI

1.3 Candidate descriptors

These descriptors will be accepted in the list of controlled terms (Lexique) if they appear often enough in the indexing in the course of one year. Candidate descriptors can be in French or German.

1.4 Free terms

By definition the free terms and the empty terms are characterized by the fact that they are not included in the list of controlled terms and that they are too general or too specific in the given context. Brand names for example belong to this category.
Free terms appear indifferently in French or in German.

1.5 Role indicators (since 1980)

All of the descriptors belonging to one of the above categories may be characterized by a role indicator. These are four in number used in this part of the PASCAL data base:

/SUB: substrate, basic materials in which a process is going on

/SEC: secondary material (or related material), deposited material, layer

/ACT: active product

/ANA: measured (determined) element in an analyzed material.

The role indicator is written directly after the descriptor to which it applies separated only by a special character as for example an exclamation mark in the IRS-QUEST system.

Ex: effect of boron on the size of steelgrains would be indexed

ACIER
BORE/ACT

The role indicators SUB and SEC are principally used in the four following cases: 1) Fiber strengthened composite materials: the matrix is adjoined the indicator SUB, the reinforcing material is marked with the indicator SEC

Ex: aluminium reinforced by tungsten fibres
ALUMINIUM/SUB or AL/SUB
TUNGSTEN/SEC or W/SEC

2) Coatings: the substratum receives the role indicator SUB, the coating material the indicator SEC

Ex: galvanized steel
ACIER/SUB
ZINC/SEC or ZN/SEC

3) Welding: the basis material is given the indicator SUB, the additional metal the indicator SEC

Ex: welding of the alloy ALZN4MG1 using the alloy ALZN3MG4 as filler wire
ALLIAGE3 AL:>50 MG:1-1,5 ZN:3-5/SUB
ALLIAGE3 AL:>50 MG:3-5 ZN:3-5/SEC

4) Diffusion: the metal in which another metal is diffused is given the indicator SUB, the diffusing metal the indicator SEC

Ex: diffusion of copper in iron
METAL PUR FE/SUB
CUIVRE/SEC or CU/SEC

In the case a descriptor needs two role indicators these follow upon each other in the most logical possible order

Ex: coating of an intermediary layer of copper on steel, followed by a coating of chrome
ACIER/SUB
CUIVRE/SEC/SUB or CU/SEC/SUB
CHROME/SEC or CR/SEC

2 CHAINING

As in the rest of the PASCAL file there exist links between the descriptors and/or the free terms: these are chains to build indexing phrases or logical word-chains. They are used principally in the published version of BS to help manual retrieval.
The logical chains (or splits) existed already in the Metallurgy part of BS but they have been extended in the frame of the Franco-German co-operation to diminish false drops in documentary searches. They are used for indexation when several different 'studies' are found in the same article, for example:
Influence of certain additions on the machinability of alloys: TE in the free cutting (copper) alloy and S in the free cutting steel.
The indexation will be:

01 ACIER DECOLLETAGE
02 ALLIAGE DECOLLETAGE
03 TELLURE/ACT
04 SOUFRE/ACT
05 USINABILITE
26 ACIER S (identifier)
27 ALLIAGE CU:>50 TE ZN (identifier)

with the following chaining:

S0 05 (S0 indicating that this term(s) belong to both chains S1 and S2)
S1 01,04,26
S2 02,03,27

meaning the following chains of terms:

USINABILITE, ACIER DECOLLETAGE, SOUFRE/ACT, ACIER S
USINABILITE, ALLIAGE DECOLLETAGE, TELLURE/ACT, ALLIAGE CU:>50 TE ZN

3 SUPPLEMENTARY CLASSIFICATION CODE 741...

May be attributed (in sect. 740) to documents specially treating materials. Can be considered as some kind of macro-indexation.

741.	Materials
741.A.01	Cast Irons and Steels
.A	Pig iron
.B	Cast and cast steels
.C	Ferrous alloys
.D	Non-alloy steels
.E	Micro-alloy steels
.F	Alloy steel
741.A.02	Nonferrous metals
.A	Nonferrous metals and alloys
.B	Nonferrous casting metals and alloys
741.A.03	Cermets
741.A.04	Non-metallic materials (used in metal processing)
741.A.05	Non-metallic by-products (ashes, slag)

Appendix

APPLICATION TO PASCAL IN THE FIELD OF METALLURGY IN THE EPOS/VIRA SYSTEM

Since 1981 PASCAL is divided in four data bases and searched as PASCMy (y=0, 1, 2, 3).

PASCM0 data base corresponds to previous PASCU8 and PASCU0. It includes the sections: 740 - METALLURGY and 745 - WELDING, BRAZING AND RELATED TECHNICS.

The indexation rules for metallurgy in PASCAL imply the following in searching PASCM0 with EPOS/VIRA system at IDC-KTHB.

PASCAL DESCRIPTOR TYPE	CORRESPONDING KOMPOST FIELD	EPOS SEARCH TYPE	SDI-CARD
French controlled descriptors	29.2	29.2 0.7 0.5	yes, at end of ref
identifiers	30.2	30.2 0.7 0.6 0.5	yes, last at end of ref
candidate descriptors	29.5	29.5 0.7 0.5	not written on output
English controlled descriptors	29.1	29.1 0.7 0.5	yes, at end of ref
German controlled descriptors	29.3	29.3 0.7 0.5	not written on output
free terms	30.2	30.2 0.7 0.6 0.5	yes, at very end of ref

These fields are searched in character mode.
Searching in character mode means that left hand truncation is allowed. It means also that even if no left hand truncation is used the searched string may be found inside of a multi-term descriptor string if it appears after a space, i.e. as it is in the case of searching in word mode.

Ex: /OXYDABLE/ (left hand truncated) will even retrieve references indexed with INOXYDABLE or ACIER INOXYDABLE

/ STEEL/ (not left hand truncated) will even retrieve references indexed for example with STAINLESS STEEL.

Identifiers are not translated to English (or German) and the French form of the prefixes has to be used for retrieval.

As there are spaces inside of the formula describing the different materials, elements or specific compounds can be retrieved each for itself and combined with use of the profile logic.
Ex: searching for alloys containg both silver and cadmium oxide

LO=A * B * C

A=30.2/ AG / (or EPOS type .5, .7, .6)
B=30.2/ CDO /
C=30.2/ ALLIAGE/ with right hand truncation if all types of alloys are acceptable.

The composition range of a specific element can be precised but all acceptable composition ranges must be searched for at least all the acceptable lower range-limits
Ex: A=30.2/ AG 50 / or A=30.2/ AG: 50 / will retrieve materials with more than 50% silver.

Ex: B=30.2/ CDO:10-20/ or B=30.2 /CDO 10 20/ will retrieve materials in which the composition range of cadmium oxide is 10-20% (and only 10-20%)

Ex: B=30.2/ CDO:10/ or B=30.2/CDO 10/ will retrieve materials in which the content of cadmium oxide has a composition range with lowest limit 10 and any upper limit

Candidate descriptors and free terms

They are found in Kompost fields 29.5 and 30.2 respectively and can be searched specifically with these EPOS types or using .5 or .7. Both French and German terms can occur in this field and both languages should be used for maximal retrieval.

Role indicators

When searching for a given substance with a given role (SUB, SEC, ACT or ANA) one has to search for all the possible positions of that role after the substance name and chemical symbol as the descriptors with their role indicators are kept in one string with the special character (a short straight line) between the descriptor and the role(s) indicator(s)

Ex: searching for silver as substractum

A=.5/ AG/SUB/	and	A=.5/ARGENT SUB/
A=.5/ AG SEC SUB/		A=.5/ARGENT SEC SUB/
A=.5/ AG ACT SUB/ etc, etc		A=.5/ARGENT ACT SUB/ etc, etc

it can be better (and easier) to search for all the references dealing with AG (silver) and combine this with the role indicator SUB searched for in the descriptor field .7

Ex: LO=A * B

A=.5/ AG / A=.5/ ARGENT/

A=.5/ SILVER/

B=.7/ SUB /

of course this will result in many false drops (where the SUB role is not related to AG) but the relevant references should be easily identifiable in the SDI output as the descriptors are written at the end of the reference on the output cards and this will at least give a fewer number of references than if one had searched on AG, silver and argent without limiting to a given role.

Classification code 741...

may be assigned to documents dealing specifically with materials. It should be used as a supplementary search term (a synonym) and not to limit the results of a search as this code is not always used to characterize relevant documents.